Chapter 3

Analyzing English Phrases from Pāṇinian Perspective

The use of computers as an information processing device demands a sound theory for processing information in a language. Pāṇinian way of analyzing a language provides such a theory. In order to support this claim, we will take some concepts from the Pāṇinian grammar and use them to analyze languages other than Sanskrit.

Morphologically, Sanskrit is a very rich language. It uses explicit morphemes to express semantic information through syntax. Now, the question is whether the concepts from the Pāṇinian grammar written for Sanskrit can be applied to languages that are not similar to Sanskrit? For example, can we model English, a language that is morphologically not so rich, using the concepts from Pāṇini’s grammar?

Sanskrit expresses agreement between adjectives and nouns or between nouns in apposition through gender, number and case markers. On the other hand, English uses the notion of ‘phrase’ to express various semantic relations among the words in a sentence. The notion of ‘phrase’ is a way to capture the semantic information through position of the elements involved in an expression. Thus, it appears that these two languages use very different mechanisms to encode the same information. In such a case, how can one analyze English sentences using the grammar written for Sanskrit?

In this chapter, we will take some Pāṇinian primitives such as pada (syntactic unit), sup (nominal inflectional suffixes), tin (verbal inflectional suffixes), etc., and model English using these primitives. We will analyze English ‘phrases’ by defining the sup and tin suffixes for English and compare them with the notions of pada and samasta-pada (compound).

3.1 Introduction

Pāṇinian grammar analyses basic syntactic unit as a pada. A pada is an inflected word form which is ready to participate in a sentence formation. Sanskrit has a grammatical rule apadaṁ na prayuñjīta [53], meaning: “a word which is not a pada should not be used in a sentence”. Another established rule is: na kevalā prakṛṭīḥ prayoktavyā na ca kevalaḥ pratayayaḥ [71] that means neither bare roots nor bare suffixes can be used in a sentence.
What does it mean? It means that a sentence expresses a thought through some words. Thus, a sentence is grammatically a whole which is made up of related words that have their individual meanings. The words by themselves do not express the meaning of a sentence. The meaning of the sentence comes from the meaning of the words and their relation to each other. The important part of how the sentence meaning is composed depends on how the individual units are related to each other. Thus, words in a sentence have two parts to them: roots/stems (content words) and grammatical markings (affixes in a morphologically rich language).

In other words, the composite meaning of a sentence is constructed through some grammatical constructions. These grammatical constructions are mutually semantically bound through some grammatical markings. We can analyse these grammatical constructions as meaningful units in terms of a combination of a “word + inflectional affix” where an inflectional affix expresses the relation of a word with other word in a sentence. Unless that grammatical marking is present, the relation is not expressed and therefore the word cannot qualify to participate in a sentence. So, for it to participate in a sentence formation, it should have appropriate marker. The notion of pada captures it, that is why only a pada will qualify for sentence formation. Without a suffix, it is only a bare root and a bare root does not carry the relational information. Its role with respect to other words has to be marked. Once we mark it, it becomes a pada.

Similarly, Sobin’s statement that “only phrases may be sentence fragments” [120] imposes similar constraint for English. This is alluring us to compare the two concepts pada of Sanskrit and ‘phrase’ of English. We investigate whether the notion of pada can be extended to English and what are the mechanisms for forming a pada in English. And, if yes, then what are the equivalent mechanisms of sup, tiṅ for forming a pada in English? Is the notion of phrase equivalent of pada or a phrase contains more than one paddas? These are the questions we will be addressing in this chapter. We will first discuss the notion of pada from Sanskrit grammar and then move on to see various English phrases, how various English phrases can be explained in terms of pada.

3.2 Related Work

Gangopadhyaya (1990) has analyzed noun phrases in Bengali and studied assignment of role and the kāraka theory following the Indian grammatical tradition. According to her: “The term phrase corresponds to the term pada in its minimal form but not in its expanded form, i.e. when a phrase is understood as a syntactic constituent consisting of more than one word.” According to her, a single word phrase corresponds to a pada but a phrase that consists of more than one word does not correspond to a pada [56]. A pada by its very definition suptiṅantam padaṁ¹ involves inflecting roots/stems with nominal and verbal suffixes. She further says that, “we do not see any term to label the intermediate syntactic unit such as phrase in traditional Indian grammar. The compound (samāsa) corresponds to the phrase in many cases.”

¹A word ending in sup (nominal suffix) or tiṅ (verbal suffix) is termed as pada in Sanskrit
Gangopadhyaya does not provide enough evidence to substantiate her conclusions. She also does not account for all phrases which consist of multiple words such as ‘brave soldiers’, ‘very intelligent boy’, etc. The question we investigate is how many sups are present in these expressions. If there is single sup, then can these also be considered as a single pada composed of multiple words?

According to Apte [5], the expression of a single idea is a word (pada) and the aggregation of two or more words without a subject or predicate is a phrase (padasamuccaya).

Apte calls a phrase as padasamuccaya, but does not give any linguistic account for it. He looks at Sanskrit from the English perspective, and therefore interprets a phrase as a group of multiple padas. But if we look at English from the Pāṇinian perspective, we find that a minimal/simple phrase corresponds to a pada or a samasta-pada. And, a complex phrase that is composed of two or more phrases corresponds to a padasamuccaya.

Bharati et al. (1995) introduced the notion of Local Word Grouping (LWG) for Indian languages such as Hindi. LWG is a notion similar to pada found in literature. In LWG, word groups are formed on the basis of local (adjacent) word information.

In his book, ‘Pāṇini: His Description of Sanskrit”, Jag Deva Singh occasionally talks about English phrases with respect to nominal compounds [116]. But since his focus was on analytical study of Aṣṭādhyāyī, he does not provide enough evidences with respect to English.

Our assessment is based on flow of information where word groups are formed on the basis of overt and covert neighboring syntactic inflections called sup and tīn for any language. This is grammatically more precise and also allows to find out syntactic elements that unite the words of a sentence into a meaningful unit.

To the best of our knowledge, we have not found any work that analyzes English phrases from the Pāṇinian perspective.

To look at English from the Pāṇinian point of view, the very first task is to find out the equivalents of sup, tīn and pada in English.

3.3 Inflectional Suffixes

Vibhakti is a general term which corresponds to ‘inflectional suffixes’ in Sanskrit. It is used for both the nominal as well as verbal inflections. In Sanskrit, the nominal inflections and the finite verb inflections are realized through suffixation. The nominal inflections express syntactic relations across words in a sentence. The verbal inflections also express one of the relation of its arguments through suffixation. However, the syntactic mechanisms for realizing the semantic relations across words in a sentence might be different in different languages. Some languages have prepositions (e.g., English, Persian etc.), some have postpositions (e.g., Hindi and other Indian languages) [14], some might have other kind of syntactic devices.
3.3.1 Nominal Suffix (sup)

In Paninian framework, the nominal inflections are termed as sup. Sup is an acronym formed from the first and the last phoneme of the list of nominal suffixes. A sup inflection carries information of number and the case marker in Sanskrit. For example, bālaka (nominal base / prātipadika) + -am (sup) becomes bālakam where the sup ‘-am’ carries the information of accusative case (dvitiyā vibhaktī) and singular number.

However, English has different sets of morphemes for marking number and case information. In English, the number information is primarily marked through a suffix ‘-s’ for plural and ‘-0’ for singular and the case is realized through prepositions or through ‘generalized vibhaktī’ (described later). For example, in ‘to the boy’ and ‘to the boys’, the preposition ‘to’ marks case information and the inflection ‘-s’ marks number information.

3.3.2 Verbal Suffix (tiṅ)

The verbal suffixes are termed as tiṅ in Sanskrit. Tiṅ is an acronym formed from the first and the last phoneme of the list of verbal suffixes. In Sanskrit, verbs take tiṅ suffixes to express tense, number, person, mode and voice. For instance, pathati (is reading) has patha (to read) as root and is inflected for present tense, first person, singular number and active voice by the tiṅ suffix ti(p).

In English, the behavior of modal and auxiliary verbs is similar to that of verbal inflections. The modal and auxiliary verbs express tense, number, etc., hence correspond to tiṅ suffixes.

3.4 Pada: A Primary Syntactic Unit

Sanskrit uses two different terms śabda and pada, both of which are roughly translated as ‘word’ in English. In Paninian system, a śabda denotes a linguistic expression ranging from individual speech sound to utterance [116]. In the sūtra (A. 8.3.86), Pāṇini uses the term śabda for the name of a particular letter (i.e. visarga) and in (A. 7.3.67) he uses it in the meaning of speech or sentence or utterance. On the other hand, the term pada is used for a primary syntactic unit that occurs in actual sentences. In other words, a śabda and a pada correspond to an ‘utterance’ and its ‘fully inflected word forms’ respectively [73]. Pāṇini defines a pada in two ways [87]:

1. A finished word form which is ready to participate in a sentence. The sūtra suptiṅantaṁ padam (A 1.4.14) states: “a unit ending in sup (nominal suffix) or tiṅ (verbal suffix) is called a pada”.

   According to this sūtra:

   (a) Prātipadika + sup = subanta pada
       (nominal stem + nominal suffix = nominal pada)

   (b) Dhātu + tiṅ = tiṅanta pada
       (verbal stem + verbal suffix = verbal pada)
2. An unfinished word form defined only for technical purpose during word formation.

The second definition of pada is specific to unfinished derivational process of Sanskrit words. The first definition talks about inflected verbal and nominal units. It takes into consideration the syntactic units of a sentence that can be extended to any natural language. In this chapter, we are concerned only with the first definition of pada. Figure 3.1 depicts pada formation process.

![Pada formation process diagram](image)

The derivation process in PG takes prātipadika (nominal stem) described in Section 3.5 and/or dhātu (verbal stem) described in Section 3.6 as basic input and terminates the process with the derivation of pada [113]. Since a pada is formed with respect to actual sentence structure, it is called the highest derivative and is a syntactic unit rather than a morphological unit in Pāṇinian system [113].

The Section 3.5 and 3.6 briefly describe the process of nominal and verbal stem formation in Sanskrit and show how padas are formed by attachment of sup and tin inflections to nominal and verbal bases respectively.

### 3.5 Prātipadika: Nominal Stem Formation

1. **Primitive**: The sūtra arthavadadhūrāpratayayah prātipadikam (A. 1.2.45) defines a primary meaningful nominal base which is neither a verbal root nor an affix nor an inflected word as prātipadika. For example, dīthā (a wooden elephant), kapitthā (a wood apple tree) etc. Gaṇapāthā, an appendix to Pāṇinī’s Āṣṭādhyāyī, gives a few lists of prātipadikas but most of the prātipadikas are derived.

2. **Derivatives**: A nominal base derived through derivational morphology is also termed as prātipadika by kṛttadhitasamāśśca (A. 1.2.46). These derivatives are mainly of three types: (a) Kṛdanta (primary derivatives); (b) Taddhītānta (secondary derivatives); (c) sāmāsa (compound). This division is based on morphological process [18, 81, 78].

   (a) **Kṛdanta (Primary Derivatives)**: The non-tin suffixes affixed to verbal stems are termed kṛt by (A. 3.1.93). Aṣṭādhyāyī describes kṛt suffixes from 3.1.93 to 3.4.76. So, the words that are derived by adding kṛt suffixes to the verbal stems are called kṛdanta. For instance, we can derive kartā (doer), kāraka (doer or agent), kṛti (creation), karana (an instrument or means of an action), karttavya (what is fit or ought to be done), karaṇiya (doable or what
is to be done) etc., by adding the kṛt suffixes -trc, -nvul, -ktin, -lyuṭ, -tavya and -anîyar respectively to the verbal stem kṛ (do).

The kṛt affixes are applied to verbal bases as well as to the verbal stems derived from verbal bases (described in Section 3.6). For example, bhoktā is derived by adding -trc affix to the verbal base bhuj (to eat) and bhojavitā is derived by adding the same affix to bhojaya (to feed) a causative derivative of bhuj.

(b) Taddhitānta (Secondary Derivatives): Taddhitāntas are derived by adding taddhita suffixes to nominal stems. The stems involved in secondary derivation should bear semantic compatibility. These constructions are optional, so the secondary derivatives (vṛtti) and canonical phrasal paraphrase [62] (henceforth abbreviated as CPP), (vigraha-vākya) both are possible. Which word takes the secondary derivative suffix is decided by the sūtra samarthānāṃ prathamādvā (A. 4.1.82). It says that a taddhita suffix is added after the first uttered word in the sūtra defining the taddhita suffix. And the padas should be samartha i.e. they must have a direct semantic relation among themselves. For example, the sūtra tasyāpātayam, segmented as tasya apatyam assigns ‘-ān’ suffix when the concerned words mean ‘his offspring’. Here the pada, tasya which is in genitive case is uttered first in the sūtra. So, the taddhita suffix is added optionally to the pada which ends in genitive case. For example, in Upagoh. apatyam, Upagoh replaces tasya because it ends in genitive case. Even if the order of the CPP (vigraha vākya), Upagoh apatyam changes as apatyam Upagoh, the taddhita suffix comes after the pada which represents the first constituent in the sūtra. It is Upagoh in our example. The taddhitānta derivative that we derive from Upagoh apatyam is Aupagava which means ‘son of Upagou’.

All nominal stems fall into these two classes in Sanskrit. The subanta padas are formed by adding sup (nominal inflections) to the primitive and derived nominal stems.

3.5.1 Samāsa (Compound) Formation

Sanskrit has a very productive compound formation system. It has four major types and 55 sub-types of compounds [83]. The word samāsa is paraphrased as samasanam samāsah which means putting together, condensing, shortening [114]. Compounding brings brevity into the expressions by dropping the nominal inflection that denotes the semantic relation among the compound members.

Two or more padas form a compound if they have direct semantic relation (śamarthya) among themselves. For example, the words in the CPP rājñāḥ puruṣaḥ (king’s man) have direct semantic relation with each other. So, they form the compound rājapuruṣa (king-man).

If the padas have a close proximity, but do not have a direct semantic relation, then compound formation is not possible. For example, same compound is not formed in bhārya rājñāḥ, puroṣo Devadattasya (wife of the king, man of Devadatta), because, in this sentence, the words rājñāḥ (‘of the king’
or ‘king’s’) and puruṣaḥ (man) lack direct semantic relation (sāmartya).

Pāṇini classifies the compounds into following four classes. The semantic function of the components of the compounds is the predominant basis behind this classification [92].

1. **Avyayībhava (Adverbial Compound):** The first component of this type of compounds is an indeclinable (avyaya) and has primacy². Being a significant term (where non-indeclinable becomes indeclinable³, this compound behaves like an indeclinable. For example, the CPP kūpasya samīpam (near the well), when compounded, becomes indeclinable upakūpam. In this compound samīpam (near), a non-indeclinable is replaced with the indeclinable or particle upa. As a result, the whole compound also becomes indeclinable (A. 1.1.40).

2. **Tatpuruṣa (Determinative):** Several varieties of syntactic patterns are found in tatpuṛuṣa compound. The first member of the compound, in its paraphrase, can end in all vibhaktis except prathamā. The meaning of the final member has primacy in tatpuṛuṣa compounds⁴. For example, a nominal stem ending in trīyā vibhakti denotes kartā (close to ‘agent’ or ‘doer’ in some cases) or karana (instrument) optionally forms a compound with a primary derivative (kṛdanta). Example, ahihataḥ derived from ahiṇā hataḥ (killed by a snake).

A tatpuṛuṣa compound is also formed with numerals, negative particles, indeclinables etc.

3. **Bahuvrīhi (Exocentric):** When two or more paddas combine with each other to denote some other person or object, the compound is called bahuvrīhi (A.2.2.23-24). It denotes an exocentric construction. It refers to an object outside its members⁵. For example, prāptam udakam yam saḥ (one who has got the water) forms the compound prāptodakah (Lit. got water), dhanus. pūn. ih. (armed with a bow) in the sense of the CPP dhanuḥ pāṇau yasya saḥ (one who holds a bow in his hand).

4. **Dvandva (Copulative):** Two or more words occurring in same vibhakti connected with each other in the sense of ca (and) form dvandva compound. The number and vibhakti the resulting nominal stem takes is that of the number of the words and vibhakti of its CPP. All the components have equal primacy in this compound⁶. For example, the nominal stem Rāmakṛṣṇa (Rama, Krishna) is derived from the paraphrase Rāmaḥ ca Kṛṣṇaḥ ca (Lit. Rama and Krishna and) which is in nominative case and contains two components, Rāmaḥ and Kṛṣṇaḥ, hence the resulting nominal stem also carries the information of dual number and nominative case as in Rāmakṛṣṇau (Rama and Krishna).

---

² Pūrvapadārthapradhānāvyayībhāvah [71].
³ anavyayam avayayam bhavati yasmin sa avayyībhāvahṣ
⁴ Uttarapadārthapradhānastpuruṣāḥ [71].
⁵ Anyapadārthapradhānō bahuvṛthih [71].
⁶ Ubbhayapadārthapradhānō dvandvah [71].
In this Section, we have talked about various ways of nominal stem formation in Sanskrit which includes derivation of nominal stem by adding primary and secondary derivative suffixes to verbal and nominal stems respectively and formation of compounds. From compound formation, one can notice that almost all types of grammatical constructions can be condensed to a single nominal form. This condensation brings brevity in the language but also reduces the flexibility of word order of the compound members by freezing their position. The prātipadikas involved in compound formation can themselves be derived prātipadikas leading to complex prātipadika formation through recursion. Figure 3.2 shows the complexity involved in nominal stem formation in Sanskrit.

![Figure 3.2: A diagram showing the complexity involved in nominal stem formation in Sanskrit](image)

The ‘dhātu/nāmadhātu’ and the ‘prātipadika’ as shown in Fig. 3.2 are the central points of nominal stem derivation. In 1, we give some examples to describe the diagram shown in Fig. 3.2.

(1) a. prātipadika + taddhita = prātipadika
tail + kan = tailaka
   ‘a little oil’ [132]

b. prātipadika + sup + taddhita = prātipadika
Upagu + rnas + an = Aupagava
   ‘son of Upagu’ [131]

c. prātipadika + sanādi2 + krt = prātipadika
pandita + kyan + lyuṭ = panditāyana
   ‘an action of becoming a scholar (from a non-scholar)’
3.6 Dhātu: Verbal Stem Formation

Similar to the prātipadikas, dhātus (verbal stems) are also of two types: (i) basic verbal stems and (ii) derived verbal stems. This section contemplates on these two.

3.6.1 Basic Verbal Stems

The basic verbal roots are listed in Dhātupāṭha. Dhātupāṭha is a supplement to Pāṇini’s Aṣṭadhyāyī. As the name suggests, it consists of exhaustive lists of primary verbal roots called dhātus along with their basic meanings in Sanskrit. So, it is a verbal root lexicon which contains around 2000 verbal roots classified into 10 gaṇas (classes) [101, 102].

3.6.2 Derived Verbal Stems

The second type of verbal stems are derived ones. These are derived by adding sanādi suffixes to various types of input stems [111]. The term sanādi refers to twelve affixes starting from san assigned by guptiṣṭhavat san (A. 3.1.5) upto niṁ by kameriniṁ (A. 3.1.30). The twelve sanādi suffixes are: 1. -san, 2. -kyac, 3. -kāmyac, 4. -kyaṁ, 5. -kyaṣ, 6. -kvip, 7. -nic, 8. -yan, 9. -yak, 10. -āya, 11. -lyan and 12. -niṁ. A word ending in a sanādi affix is termed dhātu by sanādyantā dhātavat (A. 3.1.33).

These sanādi suffixes can be further classified into three classes based on the input these are affixed to. Some of the suffixes might fall in more than one class. For instance, the affix -nic is prescribed to nominal as well as verbal stems. We will briefly discuss these with some examples in lakāra (present tense) third person singular.
1. **Sanādi1**: Some sanādi suffixes are added to the primary verbal roots. The list of sanādi suffixes consists of -san, -kyaṅ, -kvip, -yaṅ, -nic, -yak, -āya, -iyaṅ and -n. ic. For instance, the sūtra dhātoḥ karmanah samānakartrkādīcchāyāṃ vā (A. 3.1.7) optionally prescribes suffix -san. This san can be added to any primary verbal root as well as to a derived one to form desideratives. In other words, this suffix is added to a verbal stem that occurs as karma (object) of the verb ḯ (to desire) and shares the kartā (agent) with the verb ḯ. For example, paṭhitum icchati (wants to study) becomes paṭha + -san after adding -san, and after other morpho-phonemic changes, it becomes pipaṭhiṣati (he wants to study).

2. **Sanādi2**: These suffixes are attached to nominal stems. For example, the sūtra bhrśādibhyo bhuvyacverlopa`sca halah (A. 3.1.12) assigns suffix -kyaṅ to the nominal stems listed in bhrśādi class in Gaṇapatha. This suffix is assigned in the sense of ‘to become’. For example, we derive paṇḍitāyate (a non-scholar becomes a scholar) from the stem paṇḍita (scholar).

3. **Sanādi3**: This class consists of the suffixes that are optionally attached to subanta pada. During the morphological process, the sup disappears though. For example, the sūtra suṣaḥ atmanaḥ kyac (A. 3.1.8) optionally prescribes -kyac suffix to a subanta pada in the sense atmanaḥ icchati ‘to desire X (subanta) for oneself’. For instance, let us examine the construction atmanaḥ putram icchati (wants a son for oneself). In this construction after replacing atmanaḥ icchati with the suffix -kyac and deleting the it-markers, we get the derived dhātu (verbal stem) ‘putram -ya’. Since a word ending in a sanādi suffix gets the term dhātu, the internal vibhakti of karma (object) attached to putra (son) is obligatorily deleted by supo dhāruprātipadikayoh (A. 2.4.71). After morphological changes, we get the the derivative putrīya which is equivalent in syntax and semantics to their non-derivative constructions. So, the tinanta form putrīyati means ‘he wishes for a son of his own’.

Figure 3.3 gives an overview of verbal stem formation and 2 gives some examples that illustrate the Fig. 3.3.

(2) a. dhātu + sanādi1 = dhātu
   pac + nic = pāci
   ‘to make someone to cook’

b. dhātu + sanādi1 = dhātu
   pāci + san = pipācayiṣa
   ‘to desire to make someone to cook’

c. prātipadika + sanādi2 = nāmadātu
   paṇḍita + kyaṅ = paṇḍitāyā
   ‘to become a scholar (from a non-scholar)’ [130]

d. dhātu + kṛt + sup + sanādi3 = nāmadātu
   rājṛ + śtran + am + kyac = rāṣṭrīya
   ‘to desire for his/her own country’
The padas are formed by adding a vibhakti to a verbal or a nominal stem. Figure 3.4 gives a combined picture of nominal and verbal stem formation. It also shows how Sanskrit derives the syntactic entities, the padas from the primitive and derived as well as simple and complex nominal and verbal bases.

Pāṇini classifies all types of word forms in two classes: subanta and tiṇanta. Apart from the inflected verb forms (finite verbs) which are classified under tiṇanta, all other word forms fall under subanta class. The subanta class includes all nouns, pronouns, adjectives and adverbs etc., in it. Since, adverbs are indeclinables, they do not inflect for any case. That is why adverbs do not seem to carry any sup on surface. The sūtra (A.2.4.84) deletes the inflections attached to the adverbs. It suggests that at some point of time, adverbs also had been inflecting like other nominals. Hence, adverbs also fall in subanta class.

As mentioned in Section 3.5.1, compounds are formed using two or more inflected forms (saha supa (A. 2.1.4). These inflected forms always have semantic compatibility as well as syntactic relatedness amongst them. Formation of samāsa brings condensation/contraction into the language by deletion of internal sup inflections and thereby formation of more than one padas into a single prātipadika (A. 1.2.46). This leads to a definite structural order of the components and thereafter into formation of more than one padas into a single pada. This brings in brevity into the language keeping the meaning intact. Compound formation is an extremely important notion. This in a way brings in constituency in Pāṇini’s grammar. This also becomes important for us to look at the notion of phrase from Pāṇinian perspective.

### 3.7 Samartha Theory of Pāṇini and its Relation to Pada Formation

Pāṇini’s samartha theory stands as a fundamental principle for any semantic and syntactic operation. According to Pāṇini, no grammatical operation can take place, be it pada formation or sentence forma-
Figure 3.4: A diagram showing the nominal and verbal stem formation and formation of primary syntactic unit, the subanta and tiṅanta padas in Sanskrit

The word samartha is used in the following meanings [70]:

1. “Mingled together” (saṅgatārtham samarthaṁ) or
2. “Fused together” (saṃsṛṭārtham samarthaṁ) or
3. “Seen together” (samprekṣīrthāḥ samarthaḥ) or
4. “Bound together” (sambaddhārthāḥ samarthaḥ)

For a word to stand in a syntactic structure, it is necessary to pass through one of these sāmarthyas. Sanskrit grammarians further merge these sāmarthyas into two classes:

---

7The words samartha and sāmarthya are used interchangeably in Sanskrit grammar.
1. **Ekārthībhāva sāmarthya (Single Integrated Meaning):** Sometimes a grammatical operation leads to ‘unity of meaning’ [128]. In that case, the *padas* having direct semantic connection become one *pada* as in compounds and primary and secondary derivatives. This is taken care by the first two meanings of the word *samartha* “mingled together” or “fused together”. The objective of *ekārthībhāva sāmarthya* is to present compounds as one *pada* (ekapada) or as a single unit. Such as in the compositional compounds *rājapurūṣah* (king-man) derived from *rājñāḥ* *purūṣah* and *yudhiṣṭhirah* (one who is always stable in the battle) derived from *yudhi ṣṭhirah*. And also in non-compositional compounds such as *krṣṇaraspaḥ* (cobra).

In English, *ekārthībhāva sāmarthya* is seen in both compositional as well as non-compositional compounds. For example, ‘lawn tennis’, ‘bird-cage’ and ‘blackbird’.

2. **Vyapekṣā sāmarthya (Meaning-interdependence):** It says that “any grammatical operation pertaining to *pada* formation takes place if and only if the words involved in *pada* formation have direct semantic connection between them” [128]. In this case, *samartha* means “seen together” or “bound together”. For example, *kriyā* (action), *kartā* (doer), *karma* (theme/patient) etc. are seen bound together in a sentence through semantic connection. The objective of *vyapekṣā sāmarthya* is to show sentence as a single meaningful unit. *Padas* seem to carry diverse meanings but a sentence indicates a single meaning.

The example 3 from Sanskrit depicts how Pāṇini captures the flow of information through *sāmarthya* in his grammar.

(3) *Vīrāḥ* sainikāḥ *deṣaṃ* rakṣanti
   brave.PL,NOM soldier.PL,NOM country.SG,ACC protect.PR,3,PL
   ‘Brave soldiers protect the country.’

In 3, the word *rakṣanti* is a *tiṅanta pada*. It is composed of the verbal base *rakṣ* and a *tiṅ* inflection namely *-anti*. The suffix *-anti* denotes active voice, third person and plural number. A *tiṅ* inflection is assigned to a verb with respect to its compatibility (*sāmarthya*) with the doer/agent or theme/patient of the action. In active voice, the *tiṅ* suffixes express the semantic compatibility between the action and the doer/agent of the action through agreement. When the doer is expressed by a *tiṅ* suffix, the sūtra (A. 2.3.46) [129] assigns nominative case (*prathamā vibhakti*) to the doer to express nominal stem meaning (*prātipadikārtha*), gender (*liṅga*), or number (*vacana*) etc., of the doer. This also makes the nominal a *subanta pada*. In 3, the words *sainikāḥ* and *vīrāḥ* are marked with nominative case (*prathamā vibhakti*) and plural number.

Due to semantic compatibility (*sāmarthya*) (i.e. the modified and modifier relationship) between them, the words *vīrāḥ* (brave) gets a *vibhakti* similar to its modified *sainikāḥ* (soldier).

The *sāmarthya* between the words *deṣaṃ* (country) and *rakṣanti* is that of the theme/patient which is marked with the *sup*, *-am* (accusative case, singular number). It makes *deṣaṃ* (country) a *subanta pada*. 

23
From the above description, it is clear that a *pada* is a syntactic unit that takes an inflection called *vibhakti* which explicitly marks the direct semantic relation of a word with another participant. In some cases, the *vibhakti* can also be NULL (zero) but it has to be present.

In Sanskrit, a *samāsa* (compound) also termed as *samasta-pada* (compounded *pada*), two or more *padas* that have direct semantic connection between them are condensed into one word by deleting the internal *vibhakti*s. This fixes the positions of the compound members. Compounds denote the conceptual unity of meaning of its members whereby the meaning of individual members becomes less important. In order to express its semantic relation with other words a *samāsa* (compound) also takes a *vibhakti* and becomes a *samasta-pada*.

### 3.7.1 Sāmarthya and Phrase Formation

While looking at samartha theory and the concept of phrase in English, we observed that both the theories capture the same aspect of language but in slightly different ways. They both capture the coherence of words together in a well formed syntactic structure. As per English grammar, a phrase is a sequence of words or sometimes a single word that functions as a single unit within a sentence [77, 65]. The words that are closely related to each other form a syntactic constituent. As per samartha theory, for a word to become a *pada* or to form a word group with other word/s such as in compounds, it has to have direct semantic relation with the other words in the sentence. This relates to the semantic relation among *padas* in a sentence and between the members of a compound.

According to the sūtra *samarthah padavidhiḥ* (A. 2.1.1), any operation concerning to inflected word form/s (*pada/s*) i.e. *padavidhi* takes place if and only if there exists a semantic connection. So is true for constituent formation. Words in a sentence do not occur in isolation rather they form groups on the basis of their relationship with other words. Thus, the notion of *pada*, though developed for Sanskrit which has a rich inflectional and derivational morphology, can be applied to any language. If we apply it to English which is morphologically not so rich, we have to analyze English sentences from the Pāṇinian perspective, especially in terms of *sup* and *tin* and *pada* and *samasta-pada* or *ekārthībhāva sāmarthya* (single integrated meaning) and *vyapekṣā sāmarthya* (meaning-interdependence).

### 3.8 Pada in English

As mentioned before, attachment of a nominal and a verbal inflection to a nominal and verbal entities makes them *subanta* and *tinanta paddas* respectively. Since prepositions mark semantic relations of nominals in English, they correspond to the *sup vibhaktis* in English. The verbal inflections are denoted either by suffixation as in ‘jumps’, ‘jumped’ etc., or by auxiliaries as in ‘will jump’, ‘shall jump’ or by combination of both as in ‘is jumping’, ‘was jumping’, ‘will be jumping’. Therefore, a *pada* in English can be defined in following two ways:

1. Preposition + nominal stem = *subanta pada*
2. (a) Verb + verbal inflection = 埬�儺嘐
(b) Auxiliary + verb = 埬毝嘐
c) Auxiliary + verb + verbal inflection = 埬毝嘐

3.8.1 **Subanta Pada in English**

According to the above definition of *pada*, A “preposition + nominal” sequence marks a *subanta pada* in English. For instance, ‘for them’, ‘to me’, ‘in the office’, ‘according to Rama’, etc. are some examples of *subanta padas*.

In order to analyze English sentences in terms of the notion of *sup, tiṅ* and *pada*, let us take the sentences 4 and analyse it using these Pāñinian primitives.

(4) Rama gives books to Mohan.

In 4, the verb ‘gives’ has the verbal inflection ‘-es’ as a *tiṅ* suffix, so, by the definition of *pada*, we know that it is a *tiṅanta pada*.

The word ‘Mohan’ has the semantic relation of dative participant with the *tiṅanta pada* ‘gives’. The dative relation of this participant is marked through the preposition ‘to’ as a *sup*. Hence, these two words have a direct semantic connection i.e. (sāmārthya) between them. So, the preposition/sup ‘to’ and the proper noun ‘Mohan’ would form a *subanta pada*.

We also know that the *tiṅanta pada* ‘gives’ has direct semantic relation with the words ‘Rama’ and ‘books’, but, since these words do not have any overt *sup* to express their semantic connection with the *tiṅanta pada* or any other word in the sentence, how would these pass the test of being *padas*?

In order to account for the words ‘Rama’ and ‘books’ as *padas*, let us look at the notion of ‘generalized vibhakti’ described in Section 3.8.2.

3.8.2 **Generalized Vibhakti**

As was mentioned earlier, different languages have different syntactic mechanisms for marking the semantic relations. Partially free word order languages like Hindi and other Indian languages encode relations through postpositions whereas (relatively) fixed word order languages like English encode relations among words through some overt inflections or through relative positions of the words. For example, the sentences 5 and 6 have same words but the change in the position of the constituents changes the grammatical relations of them. That is why, the notion of ‘generalized vibhakti’ besides inflections was introduced by Bharati et al. (1996).

(5) Mohan invited Hari.

(6) Hari invited Mohan.
In addition to inflectional markers like prepositions, postpositions or case endings, a ‘generalized vibhakti’ also includes relative position of the participants in a sentence [10, 125]. For English, following positions play an important role as ‘generalized vibhakti’:

- **Subject Position**: The pre-verbal position, the position immediately preceding a verb marks grammatical relation of a constituent with respect to the verb [10, 3]. In 5 and 6, change in the pre-verbal and post-verbal positions of the words ‘Mohan’ and ‘Hari’ changes the grammatical relations of these words. That is because these positions also play the role of grammatical relation markers (vibhakti) in English.

- **Object Position**: The post-verbal position, the position immediately following a verb also indicates grammatical relation in a sentence by its position [10, 3]. Depending on the nature of the verb, it is the theme/patient that sits at the object position and takes accusative case, but in case of a copular verb it is the predicative adjective or nominal complement that sits at the object position and takes nominative case because it describes the subject. In 7 and 8, ‘intelligent’ and ‘a doctor’ are predicate complements which are in nominative case and have ‘object position’ as a ‘generalized vibhakti’,

  (7) Mohan is intelligent.

  (8) Mohan is a doctor.

- **TOPIC Position**: Initial positions of a clause and a sentence such as in relative clause and topicalization mark semantic relation of one of the participant. Therefore, this position also indicates a ‘generalized vibhakti’.

Thus, besides prepositions, postpositions or case endings, a ‘generalized vibhakti’ also includes relative position of the participants in a sentence. In example 4, the ‘generalized vibhakti’ is realized through subject\(^6\) position in ‘Rama’, hence, it is a *subanta pada*. Similarly, the object ‘books’ carries a ‘generalized vibhakti’ in terms of object position, hence, it is also a *subanta pada*.

If we look at the *padas* formed in sentence 4, we notice that a *pada* corresponds more or less to a ‘phrase’. Figure 3.5 shows constituency tree diagram for 4. Figure 3.6 shows the constituents marked along with the *padas* for the same sentence.

So far, we have observed that in English, not only prepositions but also relative positions of the constituents such as subject, object and TOPIC positions mark grammatical relations similar to the *vibhaktis*. These positions along with the prepositions are termed as ‘generalized vibhakti’. In order to capture the role of a ‘generalized vibhakti’ as a *sup vibhakti*, let us revise the definition for *subanta-pada* as follows:

\[ \text{‘Generalized vibhakti’ + nominal stem = subanta pada} \]

\(^6\)In linguistics, the notion of subject in ILs is much debatable [16].
In sentence 4, each phrase consists of a single lexical item, therefore, we can say that a phrase with a single lexical item corresponds to a pada from the Pāniniian perspective. These phrases hold vyapekṣā sāmarthya (meaning-interdependence) between them which confirms their semantic compatibility to participate in a sentence formation. But a phrase is not always a single word. In linguistics, a phrase is defined as follows:

**Definition of a Phrase:** A phrase is a word or a sequence of words that functions as a single unit within a clause/sentence [65, 77].

Given this definition, how would PG handle a phrase that consists of more than one word? For instance, let us take sentence 9, where all the phrases are composed of more than one word.

(9) The new students have been working on this problem.
In sentence 9, the word group ‘the new students’ is an NP and ‘have been working on this problem’ is a VP which consists of a verb group and a prepositional phrase. The verb group ‘have been working’ contains ‘work’ as a verb and ‘have been -ing’ as a tīnāauxiliaries; hence, the group ‘have been working’ together can be taken as a tīnānta pada. But, how many padas should we consider in the constituents ‘the new students’ and ‘on this problem’?

### 3.8.3 Four Characteristic Features of the Compounds (Samasta-pada)

We know that in sentence 9, linguistically the words ‘the’, ‘new’ and ‘students’ together form a close and complete unit in the constituent ‘the new students’. Similarly, the words ‘on’, ‘this’ and ‘problem’ together form a close and complete unit in the constituent ‘on this problem’. To put it in terms of PG, we would say that the members of these constituents are ‘mingled’ or ‘fused’ together to form a complete unit. From the Sanskrit grammarians’ point of view, the components of these constituents have ekārthibhāva sāmarthya. As said before, ekārthibhāva sāmarthya is found in compounds. Can we say that such phrases resemble compound constructions (samasta-padas) in Sanskrit? If yes, how does Sanskrit grammar accounts for it? For that, let us look at the characteristic properties of Sanskrit compounds and see whether such issues can be resolved.

A Sanskrit compound has following four characteristic properties:

1. **Sublopa (Elision of Internal Sup/Vibhakti)**: Elision of internal sup as opposed to an asamasta-pada (sentence) takes place in a samasta-pada. An uncompounded word group is also called vākya (sentence) in Sanskrit. Only the final element receives case inflection in compounds. For example, the genitive case marker (śaśṭhī-vibhakti) is deleted in the samasta-pada rājopuruṣaḥ (king-man), whereas in the uncompounded word group rājñāḥ puruṣaḥ (king’s man), the genitive case marker is not deleted.

2. **Avyavadhāna (No Intervention by Other Word (Pada))**: Intervention by any other word (pada) does not take place in a samasta-pada but in an uncompounded word group it can take place. For example, one can say rājñāḥ rddhasya puruṣaḥ (man of a rich king), where rājñāḥ puruṣaḥ has been intervened by rddhasya a modifier of rājñāḥ, but in a compound, rddhasya cannot modify rājñāḥ. One cannot say *rāj-a-rddhasya-puruṣaḥ.

3. **Niyatapaurvāparya (Fixed Word Order)**: Pāṇiniṇi sūtras delineate compounds using prathamā (nominative case) and trtiyā (instrumental case) vibhaktis. The padas that occur in prathamā vibhakti are termed upasárjana by (A. 1.2.43). Further (A. 2.2.30) fixes the first position for the upasárjana in a compound. In case of dvandva (copulative) and bahuvrīhi (exocentric) compounds where multiple padas fall in upasárjana category, the sūtras ranging from 2.2.31 to 2.2.38 fix their positions based on phonological, morphological, semantic and lexical categorization. Hence, the words in a samāsa occur in a fixed order. For example, the sūtra śaśṭhī (A.2.2.8) defines a subtype of tatpurusa compound. In this sūtra, the pada, śaśṭhī is in nominative case,
hence, the pada that occurs in sixth case in the paraphrasal constituent of tatpurusa compound is termed upasarjana. For instance, in rājñaḥ puruṣaḥ, paraphrase of rājapuruṣaḥ, rājñaḥ is in sixth case, so, here rājñaḥ is termed upasarjana, hence rājñaḥ occupies the first position in the compound. But in a sentence, the padas can occur freely— rājñaḥ puruṣaḥ or puruṣo rājñaḥ.

4. Aikasvarya (Accent/Stress): All the words in a samasta-pada have only one accent/stress. For example, in rājapuruṣaḥ the stress is on the final vowel/syllable, but in a sentence, rājñaḥ puruṣaḥ, both the words are stressed independently.

Having looked at the characteristic properties of samasta-padas (compounds), let us now look at the English phrases and see whether they are comparable with samasta-pada or not. Most phrases have all the four properties listed above. We will examine them one by one. For example,

All the words in the phrase ‘the new students’ have only one vibhakti that is the ‘generalized vibhakti’ in terms of subject position. Similarly, the phrase ‘on this problem’ also has only one vibhakti ‘on’ attached to both the components ‘this’ and ‘problem’. If both the words had independent vibhakti, the phrase could have looked like ‘*on this on problem’. But we see no preposition between the words ‘this’ and ‘problem’, so, there is a sublopa.

No other phrase can occur within a phrase, i.e. we cannot say ‘*the new to play students’ or ‘*on this to play problem’. So, the phrases have the property of avyavadhāna.

They have a fixed word order. One cannot say ‘new the students’ or ‘*students the new’ or ‘*students new the’. So, they follow the principle of niyatapaurvāparya.

In English, stress is on the first word in common phrases and on the noun in descriptive phrases. Table 3.1 shows stress variations in common and descriptive phrases. The words in bold have stress in these phrases.

<table>
<thead>
<tr>
<th>Common phrase</th>
<th>Descriptive phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>a sports car</td>
<td>a small car</td>
</tr>
</tbody>
</table>

Table 3.1: Showing stress variations for common and descriptive phrases.

In our example, in the phrase ‘the new students’, the stress is on ‘students’, ‘the new students’ and in ‘on this problem’, the stress is on ‘problem’, ‘on this problem’. So, there is aikasvarśya.

One might argue that English shows different stress patterns between phrases and compounds. Where phrasal expressions mostly have end-stress and compound constructions have fore-stress [104]. Therefore, one cannot consider them as an evidence for one stress (aikasvarya). The point here we want to make is that it is not that important what stress they have. What is important is that they have only one stress, either a fore-stress or an end-stress i.e. only one member of a phrase or a compound is stressed and others are left unstressed.

9 See “Learn English with Speak Method, URL: www.speakmethod.com/syllablestresssetphrases.html”
To summarize, we can say that the constituent ‘the new students’ in sentence 9 occurs at subject position, hence, as stated by Bharati et al. (1996), it carries a ‘generalized vibhakti’ in terms of subject position. Also, if the constituent ‘the new students’ occurs in a prepositional phrase, all three words take only a single preposition as in I gave a book to the new students. In sentence 9, the head noun students along with its modifiers ‘new’ and ‘the’ takes only one-vibhaktilekasup i.e. the ‘generalized vibhakti’ as subject position. Therefore, the entire word group, ‘the new students’ can be taken as a single samasta-pada. For similar reasons, ‘on this problem’ can also be taken as a samasta-pada which carries the preposition ‘on’ as a sup.

From the above observations, we can say that complex English phrases come close to compound construction in Sanskrit, except that in Sanskrit, the compound becomes a single word or a pada but in English the phrasal constituents maintain multiple word status. In other words, the phrasal constituents do not become one word. This hardly matters. Because, even the compounds constructions in English cover whole range of written styles, such as one word as in ‘milkman’ (man who delivers milk), hyphenated as in ‘milk-fever’ (disease caused by lack of the calcium contained in milk) and with white space as in ‘milk bottle’ (bottle for containing milk). Hence, unlike Sanskrit, ekapadibhāva [86] ‘becoming one pada/word’ of more than one padas/words is not a characteristic property of English compounds. In fact, in [59] and [60], Giegerich argues that the distinction between compound and phrase is neither necessary nor possible in English. Therefore, simple English phrases can be treated as padas.

So, we can say that English phrases share the properties of a compound to some extent. However, unlike Sanskrit compounds, they maintain more than one word status. Thus, although phrases share the characteristic features of a compound, they are not compounds in the strictest sense. Therefore, we name them ‘quasi-compound’ (ardhasamāsa).

What about the sentence 10? In 10, the word group ‘four things’, and ‘scientists’ are NPs; The NP ‘four things’ does not seem to have any vibhakti. In that case, how would it pass the test of being a pada? It should have some sort of vibhakti to express its relation with other participants of the sentence and form a pada. The word group ‘have been exploring about the incredibly awesome moon’ is a VP which consists of a finite verb group: ‘have been exploring’ and a prepositional phrase: ‘about the incredibly awesome moon’. The verb group ‘have been exploring’ contains ‘explore’ as a verb and ‘have been ing’ as a tīn. So, it would be treated as tīnantā pada. But what about the prepositional phrase ‘about the incredibly awesome moon’ which has a preposition/sup though, at the same time contains one more phrase in it, the adjectival phrase (ADJP) ‘incredibly awesome’? How many padas should we consider in this PP?

(10) Four things scientists have been exploring about the incredibly awesome moon.

It seems that the phrase ‘four things’ neither has any overt vibhakti nor any ‘generalized vibhakti’. Then, how would it pass the test of being a pada? Recall from the discussion in Section 3.3 that ‘TOPIC’ position also plays the role of a ‘generalized vibhakti’. In 10, the verb ‘know’ has two arguments. One argument is represented by the subject ‘scientists’. But the second argument represented by ‘four things’ is not at the object position. It has moved leftward to the sentence initial position for topicalization [65].
Thus, the NP ‘four things’ has its ‘generalized vibhakti’ in terms of TOPIC position and that is why it is a pada or to be more precise, it is a samasta-pada.

In the PP ‘about the incredibly awesome moon’, it is not the case that an external element/phrase ‘incredibly awesome’ has intervened, hence, there is a violation of one of the compounding characteristic called avyavadhāna. Rather, it is an example of embedded compounding. Here, first the words ‘incredibly’ and ‘awesome’ form a compound and then the compound ‘incredibly awesome’ forms another compound with the words ‘the’ and ‘moon’. After that, the compound ‘the incredibly awesome moon’ takes the preposition ‘about’ as a sup which makes it a samasta-pada.

To summarize, we can say that the constituent ‘scientists’ in sentence 10 occurs at subject position, hence, as stated by Bharati et al. (1996), it carries a ‘generalized vibhakti’ in terms of subject position. The NP ‘four things’ gets its sup inflection from TOPIC position. Also, if the constituent ‘four things’ occurs in a prepositional phrase, both the words take only a single preposition as in ‘I gave you money for four things’. Therefore, the whole group, ‘four things’ will be treated as a single samasta-pada. The phrase ‘about the incredibly awesome moon’ will also be taken as a samasta-pada which carries the preposition ‘about’ as a sup.

Pāṇini’s samartha theory looks for direct semantic connection among the words and assigns appropriate syntax to express the semantic connection. Similarly, the phrase structure grammar analyses the syntactic units that have direct semantic connection between them to form constituents. Thus, the notion of sāmarthya (semantic connection) is the linguistic driving force behind formation of both phrases and the padas.

Note that we are talking about the simple phrases only now. We are not talking about the complex phrases such as a VP (verb phrase) which includes a verb and zoro or more other phrases.

### 3.9 Spectrum of Flexibility in Compounds

If we look at the various compound expressions, there appears to be a continuum in terms of structural and semantic realization of the compound forms. Structurally, compounds range from expressions having internal vibhaktis to expressions where internal vibhakti is elided. For example, in Sanskrit, januṣāṇḍhah (blind from birth or born blind), kṛcchrāḷlabdhah (obtained with difficulty), vācoyuktih (appropriate speech) etc., are some of the expressions where there is no deletion (lopa) of the internal sup vibhaktis and the compound meaning is similar to that of the canonical phrasal paraphrase (vigraha-vākyā) [61].

Presence of a vibhakti expresses the relation between words. In compound expressions, internal vibhaktis and number information are not of that importance [5, 122, 128]. The relations among members of a compound are expressed through positions of the words. That is why vibhaktis are deleted. Even in expressions like kṛcchrāḷlabdhah (obtained with difficulty), apsucarah (an aquatic animal) etc., where the internal vibhakti is not deleted, we use the same expression kṛcchrāḷlabdhah, even if we wish to
say it in dual krčhrābhāṃ labdhah (obtained with two difficulties) or in plural krčrebihaḥ labdhah (obtained with many difficulties) [128].

In rājapuruṣaḥ etc., even after deletion of the internal sup, the potency of the sup takes place by pratayalakṣaṇa (A.1.1.62), that is the operations pertaining to a suffix take place even after deletion of the suffix [129]. Therefore, the component rājan is treated as a pada for morphological operations, as a result, elision of the letter n by (A. 8.2.7) can be seen on the surface. Expressions like brāhmaṇakambalāḥ (brahmin’s blanket) etc. also undergo the similar operations but no effect of the internal sup can be structurally seen on the surface.

Semantically, the expressions like rājapuruṣaḥ (king-man), brāhmaṇakambalāḥ (brahmin’s blanket) etc. are fully compositional in meaning. In such expressions, the compound meaning is similar to their components and one can construct synonymous compounds using synonyms. For instance, rājabhṛtyaḥ is the paraphrase of rājapuruṣaḥ and dvijakambalāḥ is the paraphrase of brāhmaṇakambalāḥ. On the other hand, there are expressions like krṣṇasarpaḥ which are non-compositional. These expressions have all the characteristics of compounds but irrespective of the meanings of the words involved in compounding the expressions give a more specialized meaning. In our example, the compounding words krṣṇa (black) and sarpa (snake) leave their meanings and give the special meaning ‘cobra’. Paraphrasing is also not possible in such cases. Such compounds are called nityasamāsa (completely frozen expressions). Table 3.2 gives an overview of spectrum in Sanskrit compounds. It illustrates the four afore mentioned compound properties plus ‘one word’ status and ‘paraphrasing’ with examples, where, you will notice that as one moves from left to right the degree of flexibility varies.

Table 3.2: Showing flexibility spectrum in Sanskrit expressions

<table>
<thead>
<tr>
<th></th>
<th>aluksamāsa</th>
<th>samāsa</th>
<th>nityasamāsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>sup-deletion</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>un-interruption</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>fixed order</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>one accent</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>one word</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>paraphrasing</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
</tbody>
</table>
3.9.1 Spectrum of Flexibility in English Compounds

We claim that all languages including English show flexibility spectrum in compounding. In English, the level of variation is similar to Sanskrit or even higher. We have already seen that English phrases fall under ‘quasi-compound’ class. The ‘quasi-compounds’ show the highest degree of flexibility in English.

Instances of aluksamāśa are also found in English. For example, in ‘kinsman’ (a blood relative, especially a male), marksman’ (a person who is skilled in shooting at a mark) etc., the compound members ‘kins’ and ‘marks’ are possessive forms: ‘kin’s’ and ‘mark’s’ [28]. The possessive suffix “-'s” does not vanish in compound forms.

The compounds like ‘blackbird’ denoting a bird of a particular species [28] fall under nityasamāśa class because such compounds give a very specialized meaning than the compound members.

Table 3.3 shows flexibility spectrum in English compounds.

<table>
<thead>
<tr>
<th></th>
<th>ardhasamāśa</th>
<th>aluksamāśa</th>
<th>samāśa</th>
<th>nityasamāśa</th>
</tr>
</thead>
<tbody>
<tr>
<td>a good boy</td>
<td></td>
<td>kinsman,</td>
<td>lawn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>marksman</td>
<td>tennis,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>bird-cage,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>football</td>
<td></td>
</tr>
<tr>
<td>sup-deletion</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(sublopa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>un-interruption</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(avyavadhāna)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fixed order</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(niyatapaurvāparya)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>one accent</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(alkasvarya)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>one word</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>paraphrasing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
</tbody>
</table>

What is the reason that all English phrases that consist of more than one word fall in samasta-pada class? Why cannot the components of a phrase behave as independent syntactic unit?

As we have already seen in 3, the modifier and modified relation between vīrāḥ (brave) and sainikāḥ (soldiers) is expressed by attachment of the same case marker to both the words. Similarly, if the head sainika takes some other case marker to express its relation, then, the modifier vīra also takes the same case marker such as vīrān sainikān (to brave soldiers), vīrebhyāḥ sainikebhyāḥ (for brave soldiers), vīrāgām sainikānām (of brave soldiers) etc. English marks the relation between adjectives and nouns or between two nouns through position. and attaches a single preposition to all the members of the constituent ‘brave soldiers’, ‘by brave soldiers’, ‘for brave soldiers’, ‘of brave soldiers’. As a result, the
positions of the constituent members get fixed and phrases are formed which correspond to *samastapadas*.

### 3.10 Handling Complex Phrases from Pāṇinian View Point

So far, we have accounted for simple phrases, phrases that were composed of a single word or multiple words as *padas* and *samasta-padas* respectively from PG point of view. The multiple words phrases that we looked upon were composed of a head noun and pre-modifier/s. Pre-modifier/s occur before the head noun. In contrast to pre-modifiers, nouns also have ‘post-modifiers’. Post-modifiers are phrases that occur after the head noun phrases [27, 25, 26, 40].

Similar to pre-modifiers, post-modifiers also occur equally commonly in English [27]. Some of the post-modifying phrases and clauses are replaceable with another, except that they vary in terms of the degree of explicitness they provide [104]. For instance, a relative clause post-modification is more explicit than a participle clause but a participle clause gives more explicit information than a prepositional phrase [104]. Such as the post-modifier in “the boy who is wearing the black cap” gives explicit tense information, whereas it is not so explicit in a non-finite clause, as in “the boy wearing the black cap”. But in a non-finite clause, the degree of explicitness is higher than a prepositional phrase, as in “the boy in the black cap”.

English post-modifiers are broadly classified into two classes. The post-modifiers are shown in bold letters in the examples below. They modify the noun phrases preceding them.

1. **Clausal Post-modifiers:** A clausal post-modifier is a clause that occurs after a noun phrase to modify it. These post-modifiers are of following types:

   - **Relative Clause:** Among clausal modifiers, relative clauses are the most commonly used post nominal modifiers [25]. A relative clause is a subordinate clause that modifies a noun and may start with the relative pronouns *who, whom, which, that,* etc. For example,

     (11) a. an institute **that conducts several workshops throughout the year**
     b. the boy **who speaks many languages like his first language**

   - **Ing-clause:** The ‘ing-clause’ also called ‘participle clause’ is a not-finite clause used to modify nouns. If not separated by a comma, an ‘ing-clause’ modifies the preceding noun. Example,

     (12) a. the man **wearing a blue shirt**
     b. students **attending this course**

   - **Ed-clause:** Similar to the ‘ing-clause’ an ‘ed-clause’ is also a compact way to modify preceding noun phrase.
(13)  a. courses **recommended by the professor**  
     b. classes **attended by the students**

- **To-clause**: A ‘to-clause’ is an ‘infinitive clause’ that modifies a preceding noun phrase.  
  Example,

  (14)  the first astronaut **to land on the moon**

2. **Phrasal Post-modifiers**: A post nominal phrase modifier is a noun phrase or a prepositional phrase that works as a modifier to the noun phrase preceding it. These modifiers are of following types:

- **Appositive Noun Phrase**: An appositive phrase is a modifier phrase placed immediately after the phrase it modifies. As shown in 15c, it can also come immediately before the phrase it modifies.

  (15)  a. my brother, **Rama**  
        b. my mother, **a pious lady**  
        c. The Indian prime minister, **Narendra Modi**, has huge young fan followers world wide.

- **Prepositional Phrase**: Prepositional phrases are the most commonly used post nominal modifiers [25]. Example,

  (16)  a. the students **of mathematics**  
        b. the boy **in the blue shirt**

Having looked at the structures of English complex phrases, it is time to move towards more complex sentences. It is also necessary to look at them because as mentioned earlier our primary task is to translate English into Hindi. Since these two languages follow different word order, we have to see how these units behave in each of these languages and how we can achieve better translation by using the notion of _pada_ and _samasta-pada_. The first property of compound is _sublopa_ (deletion of internal vibhaktis). It does not apply to most of the above mentioned phrases. Most of these phrases either have one or more overt _vibhaktis_ or ‘generalized vibhaktis’ internally. In order to verify this, let us look at the sentence in 17 and its constituency parse tree shown in Figure 3.7.

(17)  The man who came from Delhi wrote this poem.

By looking at the constituency tree for sentence 17, we notice that the sentence starts with the NP ‘The man who came from Delhi’. As its components, this NP contains a noun phrase ‘the man’ and a relative clause post-modifier. The relative clause in its turn contains three more phrases: a WHNP ‘who’, a verb ‘came’, and a PP ‘from Delhi’. Thus, the first noun phrase of the sentence is composed
of four different phrases. Linguistically, all these phrases represent one single unit. Therefore, all these phrases together form one single complex phrase. But, as we have said, a phrase is equal to a pada or a samasta-pada, would such complex phrases correspond to a single pada or more different padas?

If we look at the relative clause modifier in sentence 17, we find that the relative pronoun ‘who’ has a ‘generalized vibhakti’ as subject position which makes it a subanta pada, the verb ‘wrote’ is a tiñanta pada and the PP ‘from Delhi’ is a subanta pada which has ‘from’ as a sup.

Thus, by looking at the above description, we find that a complex phrase is composed of two or more padas/phrases, and all the padas maintain their ‘padaness’ (padatva) i.e. every pada/phrase involved in a complex phrase has its own vibhakti attached to it. Therefore, we cannot say that a complex phrase equals to a pada.

Now, the question is, whether we can consider such complex phrases as samasta-padas. For that, we have to see whether these phrases have the characteristic properties of a samasta-pada. Since, many of the phrases involved in a complex phrase formation have vibhaktis attached to them hence maintain their padaness and these phrasal members also maintain their independent stress patterns, we cannot call them samasta-pada either.

The name we have given to such phrases is subanta-mukhyaviśeyyaka-ekārthaka-padasamuccaya. Which means a cluster of phrases (padasamuccaya) which has a single meaning (ekārthaka) where a subanta pada is the head of that particular complex phrase (subanta-mukhyaviśeyyaka). The head phrase is named pada-pratinidhi.
The hierarchic organization (internal phrasal structure) within a sentence also supports the subanta-mukhyaviśeṣyaka-ekārthaka-padasamuccaya view.

At this point, it may be pertinent to look at Hindi structures and see how this analysis helps in correct word form generation in target language. For that, let us take the prepositional phrase “to the son of Kunti” from sentence 18.

(18) To the son of Kunti, he is the symbol of mature wisdom.

The PP ‘To the son of Kunti’ is composed of an NP ‘the son’ and a PP ‘of Kunti’. Both of these come under the scope of the preposition ‘to’. How many padas should one consider in it? Since there are two vibhaktis namely ‘to’ and ‘of’ attached to them and one vibhakti makes one pada, therefore, they should make two padas.

But, since the preposition ‘to’ has scope over the NP ‘the son’ and the PP ‘of Kunti’, the entire complex phrase should be considered as a single pada or something equivalent to the notion of pada.

In order to resolve such issues, we need some evidence that shows the effect of the scope of the preposition ‘to’ on the PP ‘of Kunti’ at the level of syntax.

If we look at the Hindi translation of 18 as shown in 19, the phrase ‘of Kunti’ is translated as ‘Kunti ke’. In this expression, the postposition ke is the oblique case form of ka. Its direct case form would be Kunti ka as shown in sentence 20. What is it that forces ka to occur in oblique case in 19 but not in 20? Or, even within the sentence 19, why does ‘ka’ becomes ‘ke’ in ‘Kunti ke bete ke liye’ but not in ‘paripakva buddhimattā kā pratīka’?

(19) Kunti ke bete ke liye vaha paripakva buddhimattā kā pratīka hai.
Kunti of son.OBL for he mature wisdom of symbol is
‘To the son of Kunti he is the symbol of mature wisdom.’

(20) Kunti kā betā paripakva buddhimattā kā pratīka hai.
Kunti of son.DIR mature wisdom of symbol is
‘The son of Kunti is a symbol of mature wisdom.’

The modifier pada, ‘Kunti kā’ (of Kunti) comes under the scope of the vibhakti of the head: ‘ke liye’ [[Kunti ke] bete ke liye] ([To the son [of Kunti]]). This morphologically affects modifier pada, ‘Kunti kā’. In order to account for such phenomenon we extended the notion of pada to pada-pratinidhi ‘representative/head pada’. In a complex phrase, it is the pada-pratinidhi, the head phrase that has direct semantic relation with other participants in the sentence; the entire subordinate phrase/s or pada/s behaves as a modifier of the head phrase. As explained so far, if the head member of a pada occurs in oblique case, then all other modifying members also take oblique case. Similarly, if the head pada has a vibhakti, all modifier padas that come under the scope of that vibhakti take oblique case. In languages like Hindi, it is reflected through syntax which allows it a greater degree of flexibility of word order as shown in 21.

(21) a. ?Bete ke lie Kunti ke vaha paripakva buddhimattā kā pratīka hai.
b. ?Bete ke lie vaha paripakva buddhimattā kā pratīka hai Kunti ke.
c. ?Vaha bete ke lie paripakva buddhimattā kā pratīka hai Kunti ke.
d. *Kunti kā bete ke lie vaha paripakva buddhimattā kā pratīka hai.
e. *Bete ke lie Kunti kā vaha paripakva buddhimattā kā pratīka hai.

In 21, the *pada, ‘Kunti kā’ (of Kunti) behaves as a modifier of ‘bete ke lie’ (to the son). The post-position ‘ke lie’ assigns oblique case to the head and all its components/modifiers including ‘Kunti kā’. As a result it has the same case that is available to the head noun betā (son). Because the semantic information is formally marked on both the padas, ‘Kunti ke’ and ‘bete ke lie’ through syntax they have got a little more freedom for word order. But, if we do not change kā into oblique case, the sentence becomes ungrammatical as reflected through 21d and 21e.

The clue for case information comes from the *subanta pada that has scope over the pada with kā vibhakti. This confirms the need of recursion in pada formation. The concept of pada-pratinidhi and subanta-mukhyaviśeyaka-ekārthaka-padasamuccaya fulfill it. Since, LWG has traditionally been defined as non-recursive, this is the major difference where the concept of pada advances the concept of LWG.

3.11 Tiṅanta pada in English

In Section 3.8, a tiṅanta pada for English was defined as follows:

Verb + verbal inflection = tiṅanta pada

OR

Auxiliary + verb = tiṅanta pada

OR

Auxiliary + verb + verbal inflection = tiṅanta pada

In this definition, auxiliary includes primary auxiliaries [104] ‘do’, ‘have’ and ‘be’ as well as modal auxiliaries [104] ‘can’, ‘may’, ‘shall’, ‘will’, ‘could’, ‘would’, etc. So, according to this definition, tiṅanta padas would include inflected verbs such as ‘gives’ ‘gave’, ‘reads’, ‘goes’, ‘grants’, ‘granted’, etc., an auxiliary plus a verb as ‘will go’, ‘can give’, ‘should read’, ‘would accomplish’ etc., or one or more auxiliaries plus an inflected verb such as ‘is going’, ‘had been going’, ‘will be done’, ‘would have given’, ‘has been giving’, etc., are examples of tiṅanta padas in English.

In literature, a quite similar concept is mentioned by the name of Local Word Grouping (LWG) for Indian languages, whereby word groups are formed based on the local (adjacent) word information [14]. In Bharati et al. (1995), adjacency of the auxiliaries and postposition is mentioned, however, these can also occur dis-contiguously, and we consider them as well.
Verbs with one or more particles are treated as multi-word verbs [104] also called phrasal verbs. We treat phrasal verbs as verbal stems. The bold faced words in the sentences in 22 are considered as tiñanta padas from PG perspective.

(22)  a. The Ganga river would be cleaned up under National Ganga River Basin Authority plan.
     b. All the guests sat down on the chairs.
     c. Krishna was brought up by Yashoda.
     d. She will check out the hostel next month.

3.12 Issues

In this section, we will take up the cases which do not pass the test of being pada or samasta-pada in a straightforward manner. This includes cases where either the vibhaktis are separated from their objects as in preposition stranding or a pada is intervened by some external elements such as floating quantifiers and adverbs.

3.12.1 Treatment of Discontiguous Elements

Adverbs of manner tell us about the way action denoted by the verbs takes place [55]. They behave as adjunct modifiers to the verbs [104]. McConnell-Ginet (1982) and Shaer (2000) claim that their interpretation differs based on their VP-internal position and VP-external position. For instance,

(23)  a. Louisa rudely answered Patricia.
     b. Louisa answered Patricia rudely. (ibid., p. 3)

Shaer (2000) argues that 23a, “can be construed as saying that Louisa’s rudeness consisted in her having answered Patricia, ... whereas 23b locates the flow in the manner of answering”.

Similarly, a sub-class of manner adverbs is called ‘modal adverb’. Modal adverbs occur with modal auxiliaries in order to add some meaning to them. Similar to manner adverbs, modal adverbs are also placed before or after the modal verbs. As briefly mentioned in Section 3.11, modal verbs are also included in auxiliaries (tiṅ inflections). Recall that English tiñanta pada may be formed in three ways: (i) as inflected verb form, (ii) one of more auxiliaries plus an inflected verb form and (iii) one of more auxiliaries plus an uninflected verb form. However, examples such as 24 show that modal adverbs can intervene the ‘auxiliary + verb’ sequence. They can occur at various positions in English as shown in 24.

(24)  a. The kids likely will have been seen.
     b. The kids will likely have been seen.
     c. The kids will have likely been seen.
Now, the question is whether we consider such adverbs as a part of tīṅanta inflections of verbs or independent entities?

Since, position of such adverbs is not fixed, we decided not to consider such adverbs as part of tīṅanta padas. So, we treat any adverbs that can be inserted between auxiliaries and verbs as distinct and independent padas and group the auxiliaries and verbs to make tīṅanta padas.

Distinction between adverbial padas and tīṅanta padas helps in correct translation of verb and auxiliaries.

In 25, the adverb ‘impatiently’ is embedded in the tīṅanta pada, ‘is waiting’. Identification of tīṅanta pada and adverbial pada distinctly has helped in correct order of verb and tīṅ inflections in Hindi. Otherwise, it would have been translated as *pratīkṣā kara besabrī se rahā hai in Hindi.

But then, since a tīṅanta pada is not a compound, other words can intervene. Even after intervention by the adverbs the phrase/sentence does not become ungrammatical due to the notion of gamakatva (the ability to convey the desired meaning). Gamakatva introduced by Patañjali in his Mahābhasya is an alternate principle to sāmarthya [134]. It accounts for the acceptable and frequent language constructions that need not follow the grammatical rules of the language strictly but convey the intended meaning without introducing any ambiguity. Since the meaning of the sentence/phrase remains intact, such interventions are allowed in language.

It should be noticed that except adverbs no other words can intervene in a tīṅanta pada. That is why, the expressions like: *is for mother waiting’, *is since morning waiting’ etc. become ungrammatical. The adverbs intervene between the finite and small set of modals/auxiliaries and the verb. These constructions are easily learnt and used by the speakers of the language. That is why, such usages are allowed in the language.

### 3.12.2 Treatment of Floating Quantifiers

The quantifiers ‘all’, ‘both’, and ‘each’ occur at more than one position in a sentence. These quantifying modifiers can float away from their head as shown in 26

(26) a. All/both the students have read the story.
   b. The students have all/both read the story.

Haegeman et al. (1999) argue that floating quantifiers originate VP internally along with its DP head. The DP moves to the subject position for case assignment. Since 26a and 26b are logically equivalent, the question is how many padas should one consider in these sentences? Do these sentences have same
number of padas or different number of padas? If we consider that the words ‘The students’ and ‘all’ form a single samasta-pada in 26b, we notice a clear violation of niyatapaurvāparya and avyavadhāna. If we say that these are two different padas, how do they both get their vibhaktis?

In Sanskrit and Hindi, the quantifiers can float, but they appear as distinct and independent padas and take same vibhakti to show case and number agreement.

(27) a. Sabhī vidyārthiyoṁ ne kahānī padhī
dl.NOM student.PL,NOM NOM story.SG read.PT,FEM
‘All the students have read the story.’
b. Vidyārthiyoṁ ne sabhī ne kahānī padhī
student.PL,NOM NOM all.NOM NOM story.SG read.PT,FEM
‘The students have all read the story.’

In 27a, the expression sabhī vidyārthiyoṁ ne is a samasta-pada (quasi-compound), whereas in 27b, the expression vidyārthiyoṁ ne sabhī ne consists of two independent subanta padas: vidyārthiyoṁ ne and sabhī ne. The same phenomenon is found in Sanskrit also. But in English, since the floated quantifiers seem to have no explicit vibhakti, how would they pass the test of being a pada?

Osborn (2013) observes that distribution of floating quantifiers is similar to that of modal adverbs. All adverbs are considered similar to indeclinables which do not inflect for any case, number etc., therefore, they do not take any vibhakti. The floating quantifiers also when float do not seem to inflect for any case or number but they seem to carry the case identical to their modifiers. Therefore, we treat the floated quantifiers as independent subanta padas. In 26a, ‘all/both the students’ is a single pada whereas in 26b, ‘the students’ is a subanta pada and ‘all/both’ is another subanta pada without any explicit vibhakti. Probably due to the principle of gamakatva (the ability to convey the desired meaning), the quantifiers are able to float in the sentence and convey the desired meaning. However, a more detailed study needs to be done to handle such issues.

3.12.3 Preposition Stranding

Preposition stranding is another phenomenon of dis-contiguous sup inflections and complements. It is a syntactic construction in which prepositions do not occur close to their complements. This phenomenon is found in many natural languages including English [66] and other Germanic and Scandinavian languages [66, 126]. For example, 28 shows some instances of stranded prepositions. In 28, the stranded prepositions are written in italic letters and their complements are written in bold and italic letters.

(28) a. Which person were you talking to?
b. Which book were you talking about?
c. This is the book, the professor was referring to.
d. This chair has been set on by the delegates.
Preposition stranding is allowed in English due to the principle of *gamakatva* (the ability to convey the desired meaning), or due to the ‘Identifiability Condition’ and the ‘More/Less Important Information Condition’ as pointed out by Takami (1992). This means that prepositions strand iff the prepositional objects can be identified by the phrases/clauses which they are placed in. The phrases/clauses containing the stranded prepositions have some expectancy that can be fulfilled only by the fronted prepositional complements.

Such syntactic constructions are seldom found in Hindi, therefore, stranded prepositions should be grouped with their objects for translation in Hindi.

3.13 Conclusion

In this chapter, we have analyzed English phrases from the Pāṇinian perspective. We have defined nominal inflection (*sup*), verbal inflections (*tiṅ*), and syntactic units like *padas* (fully inflected syntactic units) and *samasta-padas* (compounds) for English.

In order to encode the semantic information through syntax, English uses different mechanisms as *sup* (nominal inflections) and *tiṅ* (verbal inflections). *Sup* include overt morphemes like prepositions and covert mechanisms such as ‘subject’, ‘object’ and ‘topic position’. *Tiṅ* incorporate finite verb inflections like ‘-es’, ‘-ed’, ‘-en’ etc., and auxiliaries and modal verbs.

We have compared English phrases with the notion of *pada* and *samasta-pada* in Sanskrit. We have shown that a single word phrase directly corresponds to the concept of *pada* in Sanskrit. On the basis of four characteristic properties of compounds from Sanskrit, we have shown that a multi-word English phrase, a phrase consisting of more than one word, corresponds to ‘quasi-compound’ (*ardhasamāsa*) type constructions in English.

We have also accounted for complex English phrases, phrases which are composed of two or more phrases as *subanta-mukhyaviśeyaka-ekārthaka-padasamuccaya*. Which means a cluster of phrases (*padasamuccaya*) which has a single meaning (*ekārthaka*) where a *subanta pada* is the head of that particular complex phrase (*subanta-mukhyaviśeyaka*).

The research carried out in this chapter also shows that there appears to be a continuum in terms of structural and semantic realization of the compound forms in a language. Structurally, compounds range from the expressions which have internal *vibhakti* (syntactic marker) to the expressions where internal *vibhakti* is elided. Semantically, they range from the expressions which are compositional in meaning to the expressions which are non-compositional.

By analyzing English phrases from Pāṇinian perspective, the study adds one more chapter to the fact that insights from Pāṇinian Grammar can be used to analyze any language from information theoretic point of view [80, 79, 3].

In the following chapters, we will see how the concept of *sup*, *tiṅ* and *pada* can help us identify the overt and covert mechanisms of information encoding in a language. We will see how this mechanism can help us in mapping various dependency structures into one standard format and in formulation of...
transfer grammar rules in order to arrive at more natural target language word order. We will see how one can draw generalizations and reduce the number of rules subsequently.